



HEMEROTECAS DE PRENSA DIGITAL. EVOLUCIÓN Y TENDENCIAS



Javier Guallar, Ernest Abadal y Lluís Codina



Javier Guallar es profesor de la *Facultat de Biblioteconomia i Documentació* de la *Universitat de Barcelona*, de la *Facultat de Comunicació Blanquerna* de la *Universitat Ramon Llull*, y de los *Estudios de Información y Comunicación* de la *Universitat Oberta de Catalunya*. Anteriormente fue documentalista de prensa en varios diarios. Subdirector de la revista *El profesional de la información*, coordinador del *Anuario ThinkEPI* y codirector de la colección de libros *El profesional de la información* (Editorial UOC).

<http://orcid.org/0000-0002-8601-3990>

Univ. de Barcelona, Fac. de Biblioteconomia i Documentació
Melcior de Palau, 140. 08014 Barcelona
jguallar@gmail.com



Ernest Abadal es decano de la *Facultat de Biblioteconomia i Documentació* de la *Universitat de Barcelona*. Sus líneas de investigación actuales son las publicaciones digitales y el acceso abierto a la ciencia, temas sobre los que ha escrito varios artículos y monografías. Es coordinador del grupo de investigación consolidado *Cultura y contenidos digitales* y codirige un proyecto del *Plan Nacional de I+D+i* sobre acceso abierto a la ciencia en España.

<http://orcid.org/0000-0002-9151-6437>

Univ. de Barcelona, Fac. de Biblioteconomia i Documentació
Melcior de Palau, 140. 08014 Barcelona
abadal@ub.edu



Lluís Codina es profesor titular del *Depto. de Comunicación* de la *Univ. Pompeu Fabra* y director de la *Unidad de Soporte a la Calidad y a la Innovación Docente (Usquid)* de la *UPF*. Imparte docencia en los *Estudios de Periodismo y de Comunicación Audiovisual*. Fundador y codirector del *Máster en Documentación Digital* de la *UPF* y del *Anuario Hipertext.net*. Autor de unas 200 publicaciones y de 6 libros como autor o coautor de su especialidad. Coordina un proyecto sobre ciberperiodismo del *Plan Nacional de I+D+i* del *Mº de Economía y Competitividad*.

<http://orcid.org/0000-0001-7020-1631>

Univ. Pompeu Fabra, Depto. de Comunicación
Roc Boronat, 18. 08018 Barcelona
lluis.codina@upf.edu

Resumen

Se analizan las hemerotecas digitales de los diarios españoles de mayor audiencia y difusión. Es un estudio evaluativo basado en 27 indicadores divididos en 4 grandes secciones: aspectos generales, contenidos, sistema de consulta y presentación de resultados. Se muestran los resultados de un análisis realizado en 2010, y se relacionan con los obtenidos anteriormente en 2007. Se señalan los aspectos mejores y peores de las hemerotecas de prensa digital, se establece un ranking de diarios en el que ocupan las primeras posiciones *El mundo*, *Abc*, *El país* y *La vanguardia*, y se destacan 5 grandes tendencias observadas: ampliación del fondo temporal online, hemerotecas gratuitas versus hemerotecas de pago, simplificación (y empobrecimiento) de la búsqueda, filtros en la presentación de resultados y productos documentales en la sección Hemeroteca.

Palabras clave

Prensa digital, Diarios digitales, Hemerotecas digitales, Archivos de prensa, Documentación periodística, Sistemas de búsqueda, Evaluación, España.

Title: Digital press archives. Evolution and trends

Abstract

The digital press archives of the Spanish newspapers with the largest digital circulations are analysed. Evaluative study based on 27 indicators divided into 4 main sections: general features, content, query system, and presentation of results. The

Artículo recibido el 17-10-12

Aceptación definitiva: 23-11-12

analysis was conducted in 2010, and results obtained in 2007 are compared. The best and worst aspects of the digital press archives are pointed out, establishing a newspapers ranking: *El mundo*, *Abc*, *El país* and *La vanguardia* occupy the top positions. The five major trends: extension of the temporary fund online, free archives versus payment archives, simplification (and depletion) of the query, filters in the presentation of results and documentary products in the archives are described.

Keywords

Digital press, Online newspapers, Digital press archives, News librarianship, Search systems, Evaluation, Spain.

Guallar, Javier; Abadal, Ernest; Codina, Lluís. "Hemerotecas de prensa digital. Evolución y tendencias". *El profesional de la información*, 2012, noviembre-diciembre, v. 21, n. 6, pp. 595-605.

<http://dx.doi.org/10.3145/epi.2012.nov.06>

1. Introducción

Los sistemas de acceso y consulta a los contenidos retrospectivos de los diarios en internet son conocidos indistintamente con las denominaciones "archivos" o "hemerotecas" digitales. Han evolucionado desde ser en sus orígenes meros depósitos de documentos sin sistema de recuperación (y con apenas una navegación cronológica) a mostrar en la actualidad una importante variedad de formas de recuperar información, amplia cobertura temporal y diferentes posibilidades de visualización de los contenidos.

Estas prestaciones constituyen en nuestra opinión uno de los recursos de mayor interés del mundo digital desde el punto de vista de un amplio grupo de profesionales que trabajan con información periodística, como los propios periodistas, y por supuesto los documentalistas, pero también sociólogos, historiadores, politólogos e investigadores del campo cultural en general. Sin embargo, es evidente que toda la inmensa utilidad de estas memorias se desvanecería si no existiera un sistema de acceso cómodo y eficaz.

Por este motivo consideramos oportuno y necesario analizar estos servicios, que nosotros denominamos de manera preferente "hemerotecas digitales", desde una óptica documental

Del objetivo genérico señalado se desprenden tres específicos:

- Evaluar las prestaciones, características y funciones de las hemerotecas digitales de la prensa para conocer su grado de desarrollo.
- Establecer un ranking o clasificación de los diarios digitales analizados atendiendo a las prestaciones de sus hemerotecas.
- Establecer las principales tendencias en la evolución de las hemerotecas digitales y valorar su papel en un portal de prensa

Además de las referencias señaladas antes a los artículos publicados por los autores en esta línea de trabajo, cabe citar aquí los estudios que se han tenido más especialmente en cuenta: Codina (2000, 2006) sobre evaluación de recursos web; Abadal (2002) sobre evaluación de bases de datos web; Jiménez-López; González-Quesada; Fuentes (2000); Zamith (2008); López-Aguirre (2009); Rodríguez-Martínez; Codina; Redaelli-Jiménez (2010); Rubio-Lacort; Blanco-García (2010) y Palacios; Ribas (2011), sobre análisis y evaluación de hemerotecas de diarios digitales; así como Camacho; Agirreazaldegi; Ronco; Peñafiel (2010) que utiliza nuestro método para analizar diarios vascos.

La metodología de la investigación es fundamentalmente evaluativa, complementada con la revisión bibliográfica de las fuentes especializadas. Las consideraciones metodológicas sobre la elaboración de los indicadores se detallan en Guallar; Abadal, 2009; y en Guallar; Abadal; Codina, 2010. No obstante, aquí se señala que los indicadores de análisis empleados se agrupan en cuatro categorías: *Aspectos generales*, *Contenidos*, *Sistema de consulta* y *Presentación de resultados* (tabla 1); que a cada indicador se le aplica una escala basada en tres valores (el indicador es deficiente: 1; correcto pero admite mejoras: 2; excelente: 3) y que además hay un cuarto valor (0) cuando el indicador es inexistente o su implementación muy deficiente. Asimismo, a dos indicadores no se les ha asignado puntuación, incluyéndose solamente a título informativo: *Denominación* y *Coste*.

La recogida de datos se ha llevado a cabo en dos momentos distintos: entre enero y julio de 2007, cuyos resultados se publicaron en Guallar; Abadal, 2010; y entre enero y julio de 2010, que se presentan en este artículo, a la vez que se ponen en relación con los anteriores. La muestra de análisis

“No existe ninguna hemeroteca española de prensa digital con nivel alto”

Objetivos y metodología

El objetivo general de este artículo es analizar las características, prestaciones y funcionalidades de las hemerotecas de la prensa digital a partir de un modelo basado en indicadores aplicado a los diarios españoles de mayor audiencia y difusión. El artículo se inscribe en una línea de trabajo sobre análisis y evaluación de hemerotecas digitales, como fruto de la cual se ha elaborado un método de análisis y evaluación de las mismas, revisado sucesivamente desde 2007 (Guallar; Abadal, 2009; Guallar; Abadal; Codina, 2012); aplicado a muestras de diarios de Catalunya (Abadal; Guallar, 2008) y de España (Guallar; Abadal, 2010).

Apartado	Indicador
A. Aspectos generales	A.1. Denominación
	A.2. Ubicación del acceso a la hemeroteca
	A.3. Información sobre la hemeroteca
	A.4. Coste
	A.5. Productos documentales
B. Contenidos	B.1. Colecciones disponibles
	B.2. Formatos disponibles
	B.3. Cobertura temporal
C. Sistema de consulta	C.1. Tipos de consulta por interrogación (palabras clave)
	C.2. Lenguaje de consulta
	C.3. Visualización de los índices
	C.4. Ajuste de relevancia
	C.5. Acotación temporal
	C.6. Acotación en colecciones
	C.7. Acotación en formatos
	C.8. Acotación por campos del registro
	C.9. Reutilización de búsquedas
	C.10. Consulta de tesauros u otros vocabularios controlados
	C.11. Acceso por navegación cronológica (browsing)
	C.12. Ayudas
D. Presentación de resultados	D.1. Gestión de las listas de resultados
	D.2. Campos visualizados del registro del documento
	D.3. Agrupación de resultados por categorías
	D.4. Identificación de los términos de búsqueda en el documento
	D.5. Formatos de visualización de un mismo documento.
	D.6. Gestión de los documentos recuperados
	D.7. Presentación de documentos relacionados

Tabla 1. Indicadores para la evaluación de hemerotecas

mayoría de diarios digitales estudiados al más alto nivel. La *Gestión de los documentos recuperados*, con variadas posibilidades de difusión en redes sociales, y la *Ubicación del acceso a la hemeroteca* en lugares bien visibles y destacados de la página principal del sitio web son los que presentan valores más altos, y están presentes por tanto de manera generalizada en todos los medios.

Buena parte de los indicadores considerados se sitúan en niveles medios de implantación. Es significativo que, dentro de éstos, los del apartado *Contenidos* alcanzan una buena posición. Se puede decir por ello que los contenidos disponibles en las hemerotecas de prensa digital española en 2010 tienen un nivel medio-alto. Es visible un notable avance en este sentido con la presencia de importantes fondos documentales en los sitios web de los diarios, variedad

de colecciones y aumento de contenidos multimedia (sobre esto último, véase Guallar; Rovira; Ruiz, 2010).

El apartado de *Presentación de resultados* muestra asimismo valores medios en los cybermedios estudiados, y llamativamente, no hay indicadores con valores bajos. Hay niveles aceptables por tanto en aspectos como *Campos visualizados del registro del documento*, *Presentación de documentos relacionados*, o las opciones de *Gestión de las listas de resultados*.

En este artículo se presentan de manera sintética los resultados obtenidos, y se discuten con detalle los principales hallazgos de la investigación desglosados en tres apartados, que responden a los tres objetivos específicos señalados antes: prestaciones de las hemerotecas digitales, ranking de diarios españoles y grandes tendencias en la evolución de las hemerotecas.

3. Resultados

Se presenta el resumen de los resultados obtenidos [1] en el análisis de hemerotecas de diarios en 2010 en la tabla 3 y el gráfico 1, y se discuten con detalle, relacionándolos con los obtenidos en 2007, en los apartados siguientes.

4. Prestaciones de las hemerotecas

El análisis se presenta en dos grupos: características mejor situadas y que se pueden considerar aceptables; y negativas y que por ello deberían mejorar, junto a algunas recomendaciones.

4.1. Aspectos aceptables

Uno de los resultados más llamativos es que muy pocos indicadores alcanzan valores altos. Es decir, hay pocos indicadores que estén fuertemente generalizados en todos o en la

En conjunto se aprecian mejores prestaciones en las páginas de resultados que en las posibilidades de búsquedas. Significativamente, en los indicadores del *Sistema de consulta*,

Diario	Url hemeroteca
Abc	http://www.abc.es/hemeroteca
As	http://www.as.com/busca/buscando.html
El mundo	http://www.elmundo.es/hemeroteca
El país	http://www.elpais.com/archivo/buscador.html
El periódico	http://www.elperiodico.com/info/archivo
La vanguardia	http://www.lavanguardia.es/hemeroteca
La verdad	http://www.laverdad.es/murcia/hemeroteca
Libertad digital	http://www.libertaddigital.com/servicios/buscador
Marca	http://www.marca.com
Mundo deportivo	http://www.elmundodeportivo.es
Periodista digital	http://www.periodistadigital.com
Sport	http://www.sport.es/buscador.asp
20 minutos	http://www.20minutos.es/archivo

Tabla 2. Relación de medios analizados, con las direcciones de sus hemerotecas

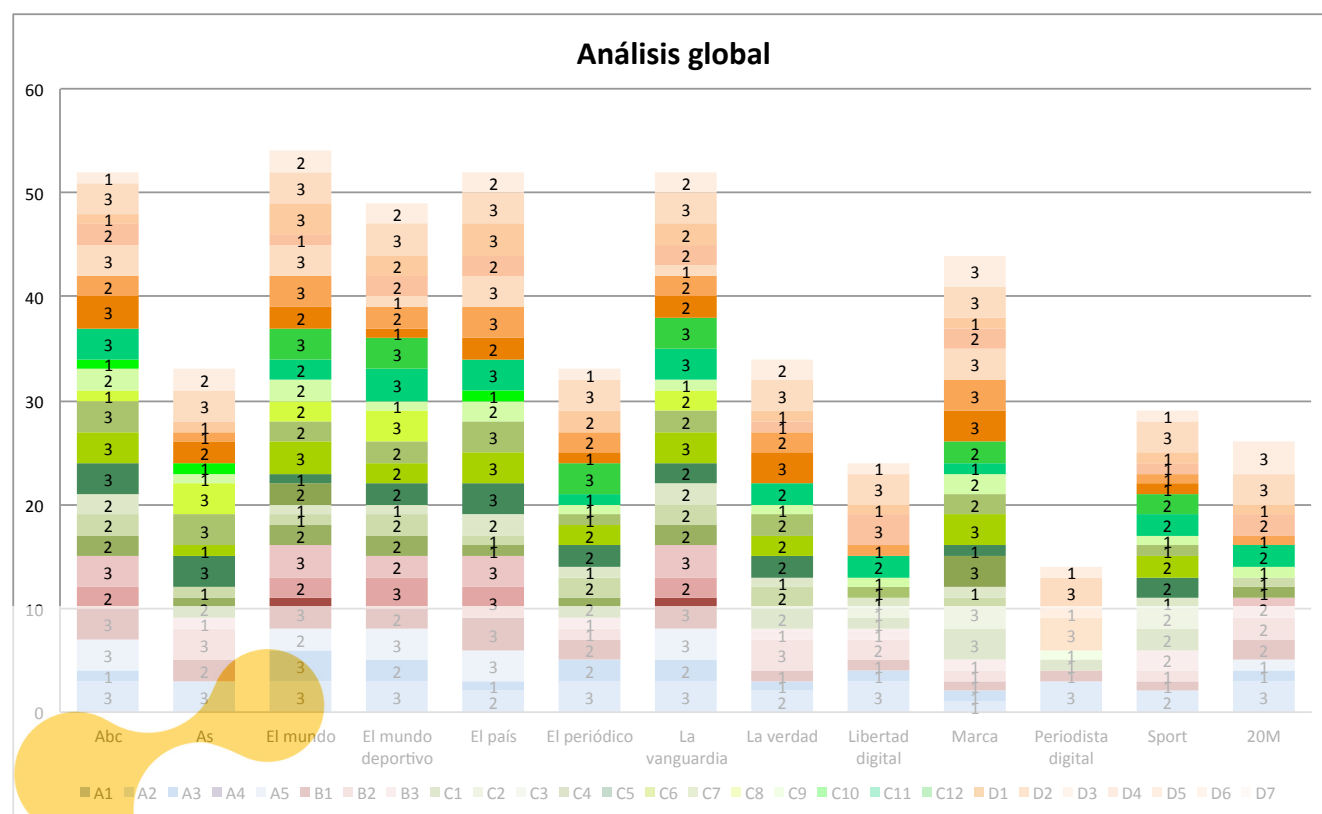


Gráfico 1. Análisis global de hemerotecas de diarios españoles

SCIPEDIA

el mejor valor (medio-alto) corresponde a la *Consulta por navegación cronológica (browsing)*, una prestación veterana altamente extendida entre la prensa digital. No es casual que se trate del único indicador de este grupo que no está relacionado con alguna funcionalidad de consulta por interrogación o palabras clave.

Algunas funciones de búsqueda parametrizada, como las opciones de *Acotación por colecciones y por formatos*, son mayoría los elementos analizados que permanecen situados a un nivel medio-bajo o bajo.

Register for free at <http://www.scipedia.com> to download the version without the watermark

Asimismo, es interesante señalar la aparición de dos interesantes servicios un poco antes de 2010 en algunos diarios (aunque presenten debido a su novedad valores medio-bajos en el conjunto de la muestra). Se trata de la *Agrupación de resultados por categorías* y la existencia en la sección de *Hemeroteca de Productos o servicios documentales* relacionados con la misma, más allá de las opciones de consulta de la misma, sobre los que se volverá más adelante en los apartados 6.4 y 6.5. En ambos casos sería muy deseable su extensión a más cabeceras de prensa digital.

Los dos indicadores con peores resultados corresponden a los relacionados con la evaluación de la calidad de la información especializada, como la posibilidad de *Ajuste de relevancia* en las búsquedas o la *Consulta de tesauros u otros lenguajes documentales*. También es necesario llamar la atención sobre el nivel medio-bajo de los dos indicadores relacionados con aspectos informativos al usuario: la *Información sobre la hemeroteca* (sobre contenidos, colecciones, opciones de consulta...) y la *Ayuda al sistema de consulta*. La información a los usuarios es algo a lo que tradicionalmente las hemerotecas de la prensa han prestado escasa atención, y esta situación no ha cambiado de manera apreciable.

A partir de lo señalado, se considera que las hemerotecas de prensa digital deberían prestar especial empeño en mejorar en los siguientes aspectos:

- generalización de las posibilidades de lenguaje de consulta profesional con la plena incorporación de los operadores de búsqueda, algo que se considera fundamental;
- generalización de las opciones de búsqueda parametrizada por campos del registro, formatos, colecciones y temporal;
- mayor implantación que la actual de algunas funciones avanzadas de consulta como la reutilización de búsquedas y la visualización de índices, así como una mayor extensión de las antes señaladas con los valores más bajos (ajuste de relevancia y consulta de tesauros) y que tienen

“ Las mejores hemerotecas digitales de la prensa española son *El mundo, Abc, El país y La vanguardia* ”

4.2. Aspectos mejorables

La consulta por interrogación, clave de todo buen sistema de búsquedas, todavía no alcanza en 2010, casi veinte años después de la aparición de la prensa en internet, los valores medio-altos o altos que todo profesional desearía. Aunque en el apartado del Sistema de consulta logren valores acep-

Indicadores	Abc	As	El mundo	Mundo deport.	El país	El per.	La vang.	La verdad	Lib. dig.	Marca	Per. dig.	Sport	20 min.	Total
A.1. Denominación														0
A.2. Ubicación del acceso a la hemeroteca	3	3	3	3	2	3	3	2	3	1	3	2	3	34
A.3. Información sobre la hemeroteca	1	0	3	2	1	2	2	1	1	1	0	0	1	15
A.4. Coste														0
A.5. Productos documentales	3	0	2	3	3	0	3	0	0	0	0	0	1	15
Total A	7	3	8	8	6	5	8	3	4	2	3	2	5	
B.1. Colecciones disponibles	3	2	3	2	3	2	3	1	1	1	1	1	2	25
B.2. Formatos disponibles	2	3	2	3	3	1	2	3	2	1	0	1	2	25
B.3. Cobertura temporal	3	1	3	2	3	1	3	1	1	1	0	2	2	23
Total B	8	6	8	7	9	4	8	5	4	3	1	4	6	
C.1. Tipos de consulta por interrogación	2	2	2	2	1	2	2	2	1	3	1	2	1	23
C.2. Lenguaje de consulta	2	1	1	2	1	2	2	2	1	3	0	2	1	20
C.3. Visualización índices	2	0	1	1	2	1	2	1	1	1	0	1	0	13
C.4. Ajuste de relevancia	0	0	2	0	0	0	0	0	0	3	0	0	0	5
C.5. Acotación temporal	3	3	1	2	3	2	2	2	0	1	0	2	0	21
C.6. Acotación en colecciones	3	1	3	2	3	2	2	2	0	3	0	2	0	23
C.7. Acotación en formatos	3	3	2	2	3	1	2	3	1	2	0	1	0	23
C.8. Acotación por campos registro	1	3	2	3	0	0	2	0	0	0	0	0	0	11
C.9. Reutilización de búsquedas	2	1	2	1	2	1	1	1	1	2	1	1	1	17
C.10. Consulta de tesauros / otros vocabularios	1	1	0	0	1	0	0	0	0	0	0	0	1	4
C.11. Acceso por navegación cronológica	3	0	2	3	3	1	3	2	2	1	0	2	2	24
C.12. Ayudas	0	0	3	3	0	3	3	0	0	2	0	2	0	16
Total C	22	15	21	21	19	15	21	15	7	21	2	15	6	
D.1. Gestión de las listas de resultados	3	2	2	1	2	1	2	3	0	3	0	1	0	20
D.2. Campos de visualización por registro	2	3	3	2	3	2	2	2	1	3	3	1	1	28
D.3. Agrupación de resultados por categorías	3	0	3	1	3	0	1	0	0	3	0	0	0	14
D.4. Identificación términos de búsqueda en el documento	2	0	1	2	2	0	2	1	3	2	0	1	2	18
D.5. Formatos de visualización de un documento	1	1	3	2	3	2	2	1	1	1	1	1	1	20
D.6. Gestión de los documentos recuperados	3	3	3	3	3	3	3	3	3	3	3	3	3	39
D.7. Presentación de documentos relacionados	1	2	2	2	2	1	2	2	1	3	1	1	3	23
Total D	15	9	17	13	18	9	14	12	9	18	8	8	10	
Total	52	33	54	49	52	33	51	35	24	44	14	29	27	

Tabla 3. Análisis global

una presencia irrelevante en estos momentos;

- mejora sustancial de las informaciones que se ofrecen a los usuarios, tanto de los sistemas de ayuda a la búsqueda, como de la información en general sobre la propia hemeroteca.

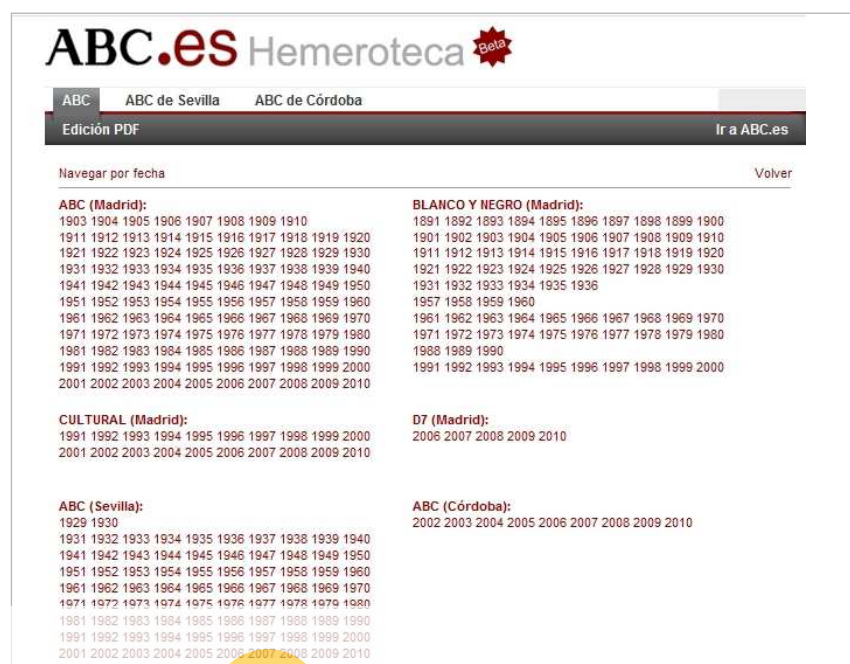
5. Ranking de hemerotecas de diarios españoles

Se ha identificado la situación en la que se encuentran los medios analizados en cuanto a la calidad de sus hemerotecas digitales. El ranking resultante se muestra en la tabla 4.

5.1. Nivel medio-alto

Si se establecen cuatro niveles a modo de cuartiles, ningún diario alcanza a situarse en la zona superior que correspondería al primero. No existe por tanto ninguna hemeroteca de la prensa digital española con un nivel alto, que pueda ser comparable al de los mejores servicios profesionales de consulta de información periodística, como las bases de datos de prensa *Lexis Nexis* o *Factiva*.

Los diarios mejor clasificados se encuadran en el segundo

Figura 1. Amplio fondo temporal de la hemeroteca de *Abc*

nivel establecido, el medio-alto. Las cuatro primeras hemerotecas son por este orden *El mundo*, *Abc*, *El país* y *La vanguardia*, con diferencias mínimas de puntuación entre ellas y que por ello son poco significativas. Estos cuatro cibermedios eran asimismo los primeros en el ranking resultante del estudio realizado en 2007 (Guallar; Abadal, 2010), aunque en otro orden (*La vanguardia*, *Abc*, *El país* y *El mundo*). Son por tanto, las mejores hemerotecas digitales de la prensa española en el período estudiado (2007-2010).

El mundo, tras el inicio en 2010 de un nuevo sistema de acceso para suscriptores (*Orbyt*), ha rebajado algo las prestaciones del acceso gratuito web mientras tiende a ir incrementando los nuevos servicios de archivo en *Orbyt*. A pesar de algunos desajustes, el equilibrio entre ambos sistemas junto a las posiciones ya altas de partida al inicio del período de análisis, especialmente en las prestaciones de búsqueda (con funciones avanzadas como el ajuste de relevancia) y su tradicional preocupación por aspectos como las ayudas e informaciones, explican esta posición.

Abc, con un fondo centenario, opta por incorporarlo de manera gratuita a su sitio web (exceptuando los últimos días) siguiendo el modelo de *La vanguardia*, diario con el que comparte varias características de orientación y estrategia. La importancia concedida por la empresa periodística a su hemeroteca digital se traduce en ese nivel medio-alto alcanzado por el conjunto de sus prestaciones, destacando además en la rápida incorporación de novedades, como los productos documentales en la sección de hemeroteca.

El país lleva a cabo en el período estudiado un cambio importante de enfoque en su hemeroteca digital, que le hace retroceder en algunas prestaciones respecto a su mejor momento, que se puede situar en 2009 (Guallar; Abadal, 2009), para quedar sin embargo en una posición en el ranking similar a la que tenía en 2007. Los cambios son tanto en el sistema de consulta como en la página de resultados: por

una parte suprime la búsqueda avanzada, elimina opciones de booleanos y algunas acotaciones, mientras que por otra parte mejora su página de resultados siendo el primer diario español en incorporar filtros. Cambio de modelo, de una búsqueda más tradicional a otra más novedosa, aunque en el balance final salgan perdiendo algunas prestaciones puramente documentales.

La vanguardia es la primera de las tres grandes hemerotecas centenarias de la prensa española que optó por situar todo su archivo en la Web y dejarlo en su mayor parte de acceso gratuito, marcando una línea que después siguieron *Mundo deportivo* y *Abc*. A pesar de alguna ligera confusión en 2010 entre los fondos accesibles vía buscador, edición impresa y hemeroteca centenaria, la relevancia que ejerce esta última en el conjunto de la prensa digital española es evidente, no solamente por lo dicho

sino también por su capacidad para marcar nuevas líneas como la inclusión de productos y servicios documentales en la sección de Hemeroteca.

Algo detrás, es meritoria también la posición de *Mundo Deportivo* y de *Marca*, que estaban además en lugares bastante inferiores en 2007: el primero ha conocido un gran impulso gracias a la incorporación a la web de su hemeroteca centenaria, mientras que *Marca*, que carecía por completo en 2007 de sistema de búsqueda por interrogación, da un salto espectacular en 2010 con unos niveles muy loables particularmente en el sistema de consulta.

5.2. Niveles medio-bajo y bajo

La verdad, *As*, *El periódico*, *Sport*, *20 minutos* y *Libertad digital* tienen un nivel de desarrollo de sus servicios de hemerotecas que se puede considerar medio-bajo, y quedan con una diferencia importante respecto al grupo anterior de diarios. En último lugar queda, rezagado y con un nivel de prestaciones bajo, *Periodista digital*.

1. <i>El mundo</i>	54
2. <i>Abc</i> , <i>El país</i>	52
4. <i>La vanguardia</i>	51
5. <i>Mundo deportivo</i>	49
6. <i>Marca</i>	44
7. <i>La verdad</i>	35
8. <i>El periódico</i>	33
9. <i>As</i>	33
10. <i>Sport</i>	29
11. <i>20 minutos</i>	27
12. <i>Libertad digital</i>	24
13. <i>Periodista digital</i>	14

Tabla 3. Ranking de hemerotecas de diarios

De este grupo se debe comentar la situación de *El periódico*, que tenía mejor posición en 2007. En 2010 se aprecia que se ha estancado e incluso retrocedido en algunas prestaciones. En su caso las enormes diferencias entre el correcto sistema profesional y de pago de búsqueda externalizado a *Mynews* por una parte, y el deficiente buscador de la edición digital por otra, explican esa posición en el ranking.

La permanencia en los puestos inferiores durante los años estudiados del diario gratuito y de los dos nativos digitales señala una situación que parece difícil de cambiar. Ello muestra que existe una distancia importante entre las ediciones digitales de las cabeceras tradicionales de prensa impresa y estos diarios, y que esta distancia no aminora entre 2007 y 2010 sino que incluso en algún caso aumenta (*Periodista digital*).

6. Tendencias en la evolución de las hemerotecas digitales

A partir del análisis de los sistemas de hemeroteca de los diarios españoles de mayor audiencia y difusión en el período 2007-2010, podemos destacar algunas tendencias generales identificadas en la evolución de hemerotecas de prensa digital.

6.1. Ampliación del fondo temporal online

Una de las tendencias más notables es el aumento de la cobertura temporal de las hemerotecas. Ya no es concebible un diario digital de calidad con un fondo documental escaso en su sitio web. Es un avance destacado respecto a la situación durante los primeros años de vida de los cybermedios.

En los mejores ejemplos, se ha llevado a cabo el escaneo y la puesta a disposición de los usuarios del conjunto del fondo documental del diario. En algunos casos se puede tratar de fondos documentales que superan los 100 años de antigüedad, como en *Abc*, *Mundo deportivo* y *La vanguardia*. Estas hemerotecas centenarias se ofrecen en pdf de manera muy

visible en el sitio web y como marca de calidad de la cabecera periodística.

Es un fenómeno remarcable: el archivo del diario de papel multiplica su relevancia a través precisamente de las facilidades que permite su acceso online: la simbiosis impreso-digital que este hecho supone es muy destacable y creemos que aún no ha sido suficientemente valorada.

“Existe una distancia importante entre las hemerotecas digitales de las cabeceras tradicionales de prensa impresa y las de diarios nativos digitales”

6.2. Hemerotecas gratis versus de pago

Década y media después de la aparición de los primeros diarios en la Web subsisten las dudas y las diferencias de criterio sobre la conveniencia de dejar su acceso libre o hacer pagar por acceder a su contenido. Las hemerotecas digitales de los diarios entran de lleno en este problema, ya que el fondo documental del diario es en principio uno de los servicios de valor añadido de la prensa digital que pueden ser más fácilmente susceptibles de comercializar. A priori al menos, porque la práctica habitual de consumo de información digital en la Web ha hecho a los usuarios generalistas (exceptuando la situación profesional) preferir con mucha diferencia el acceso gratis al de pago.

En lo que se refiere al acceso a las hemerotecas, los diarios se organizan en torno a uno de los dos grandes modelos existentes, el gratuito o el mixto (combinación de pago + acceso gratis). La tendencia en los últimos años es hacia la mayor extensión de este último, acentuada sin duda por la crisis de modelo de negocio de la prensa escrita y su búsqueda por mejorar sus ingresos.

Register for free at <https://www.scipedia.com> to download the version without the watermark

Ayuda

Tutorial



Descubre ORBYT.

EL MUNDO en ORBYT. es una aplicación que no sólo permite leer el periódico a través de internet, sino que supone reinventar la manera de entender la información. Además de leer en cualquier lugar del mundo, a cualquier hora y en cualquier soporte la edición en papel de EL MUNDO y todos los suplementos y revistas de la semana, ofrece la posibilidad de consultar la hemeroteca, opinar, interactuar con la redacción o participar activamente en el día a día de esta cabecera a través de las cuatro grandes áreas que lo componen: Quiosco, Tu Mundo, Documenta y Dutyfree.

Para acceder a ORBYT sólo es necesario teclear la dirección www.orbyt.es o hacer clic en el acceso directo que habrá siempre en la portada de ELMUNDO.es, donde contará con un espacio propio que se irá renovando constantemente. Orbyt estará disponible a un precio más que competitivo: 14,99 euros al mes o 0,60 euros al día. Cada abonado contará con una clave que le permitirá acceder desde tres ordenadores distintos, tantas veces como crea conveniente. Por ejemplo, podrá acceder desde el ordenador de casa, desde el del trabajo, el portátil y también desde el iPhone.

Figura 2. Un ejemplo de hemeroteca de pago es Orbyt de *El mundo*

El sistema gratuito es seguido de manera generalizada por los diarios nativos digitales y obviamente por los que ya son gratuitos en la prensa impresa (casos de *20 minutos*, *Liber-tad digital* y *Periodista digital*), pero es interesante señalar que también optan por esta modalidad algunos medios tradicionales de prensa escrita, como los deportivos *As* y *Mundo deportivo*. Encontramos pues dos extremos: aquellas cabeceras que muestran la hemeroteca gratuita como parte de su estrategia de presencia en la Web (*Mundo deportivo*) y aquellas con pocos contenidos retrospectivos y escasas prestaciones (*Periodista digital*), ya que no se puede hacer pagar por un servicio de poca calidad.

Por su parte el modelo mixto, es decir la combinación de contenidos gratuitos y de pago, presenta a su vez una diversidad de fórmulas. Entre ellas se pueden apreciar dos grandes orientaciones: la tendencia a ofrecer amplios fondos gratuitos, combinados con servicios de actualidad de pago (el modelo que siguen *La vanguardia* y *Abc*) frente a la de potenciar más los servicios de pago, hacia los que se dirigen las mejores prestaciones de hemeroteca (caso de *El mundo* y *El periódico*). En el primer modelo, la parte sustancial del archivo es gratuita. En el segundo, es de pago.

6.3. Simplificación (y empobrecimiento) de la búsqueda

Como los profesionales de la documentación bien saben, los mejores sistemas de búsqueda por interrogación suelen presentar dos formas: la *búsqueda simple*, para la gran mayoría de los usuarios, consistente en una única casilla en

la que escribir los términos de la consulta (el modelo *Google*); y la *avanzada*, con la posibilidad de una utilización más amplia de lenguaje documental (usualmente, operadores booleanos enmascarados tras las expresiones, “con todas las palabras”, “con alguna de las palabras”, etc.), y de parametrizar (o acotar) la consulta. Incluso en algunos casos, se puede encontrar una tercera opción denominada *búsqueda profesional*, en la que una única casilla permite a los usuarios más avanzados escribir ecuaciones de búsqueda complejas.

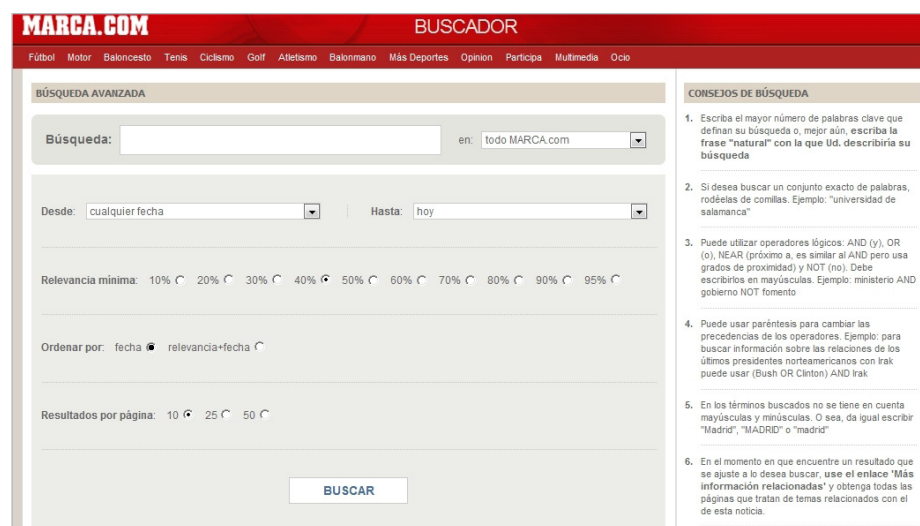
“Ya no es concebible un diario digital de calidad con un fondo documental escaso en su sitio web”

Dejando de lado esta tercera opción, que en los diarios ya no se ha llegado a implementar, en la evolución de los sistemas de hemeroteca de la prensa digital se está tendiendo de manera preocupante a nuestro juicio en algunos medios hacia una excesiva simplificación de la consulta, lo que está provocando un empobrecimiento de estos sistemas desde el punto de vista documental.

En efecto, tras unos primeros años de la prensa digital en que los sistemas de búsqueda fueron mejorando de manera paulatina, en algunos casos se ha producido un cierto estancamiento y retroceso. Algunos cibermedios que pasaron de tener solamente una búsqueda sencilla a ofrecer las dos posibilidades (sencilla y avanzada) han retrocedido y vuelven a ofrecer de nuevo solamente una única búsqueda simple. Las recomendaciones de *gurus* de la usabilidad como *Nielsen* en contra de la búsqueda avanzada (“search as simple as possible”), parecen haber sido seguidas al pie de la letra por determinados medios. Un caso paradigmático de esto es el ya señalado antes del diario *El país*, que tras alcanzar las mejores prestaciones de búsqueda de su historia en 2009, opta por volver atrás en 2010, eliminar la búsqueda avanzada y con ella varias de sus opciones de acotaciones. Con otras variantes, esta simplificación y empobrecimiento también se dan por ejemplo en los buscadores de las ediciones digitales (gratuitas) en *El mundo* o en *El periódico*.

Esta tendencia a la simplificación está produciendo en los medios que la llevan a cabo un

Register for free at <https://www.scipedia.com> to download the version without the watermark



Figuras 3 y 4. Contraste entre la única búsqueda simple de *El país* y la avanzada en *Marca*

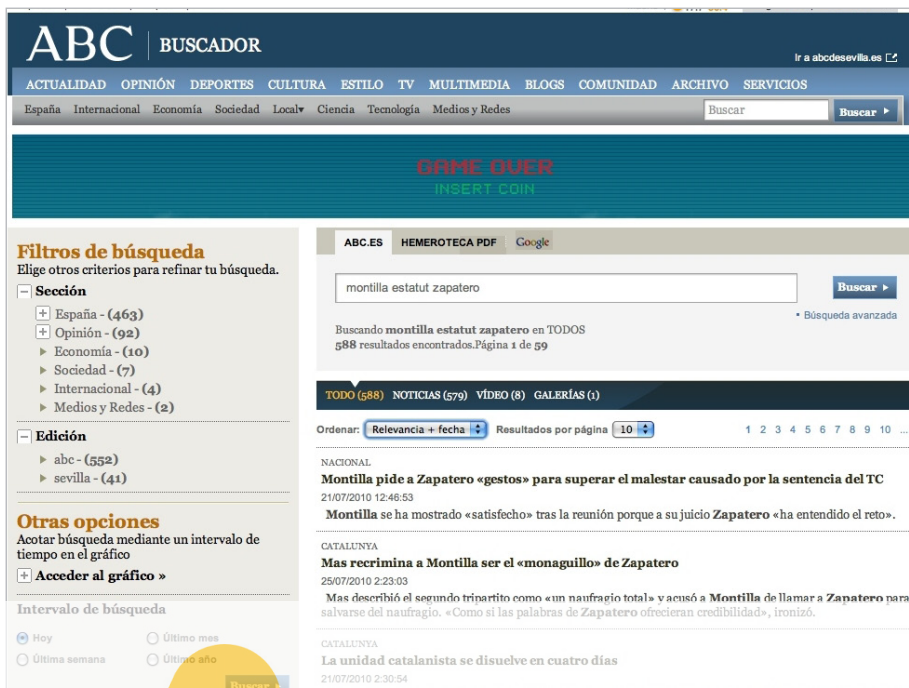


Figura 5. Filtros en la página de resultados de *Abc*

empobrecimiento de los sistemas de consulta, al limitar algunas prestaciones que se pueden considerar necesarias para garantizar la calidad de la misma, como una diversidad de opciones de búsqueda parametrizada (por campos del registro, temporal, etc.) y un lenguaje documental booleano completo. Además, algunas prestaciones más avanzadas, como la consulta de índices, reutilización de búsquedas, ajustes de relevancia... han quedado muy reducidos a pocos diarios y están muy lejos de poder extenderse entre la prensa digital española.

Register for free at <https://www.scipedia.com> to download the version without the watermark

6.4. Filtros en la presentación de resultados

Las páginas de resultados de las hemerotecas de los diarios se hacen más atractivas para los lectores y usuarios del medio mediante la incorporación de filtros (como sección, fechas, etc.) junto a los resultados de la consulta, usualmente en una columna vertical a la izquierda de los mismos. Es lo que hemos denominado *Agrupación de resultados por categorías*.

Ésta ha sido una prestación inexistente en las hemerotecas en 2007 y que algunos diarios españoles empiezan a implementar a partir de 2010, tras *El país*, el primero en hacerlo. Algunas bases de datos comerciales, como es el caso de *Factiva* en el sector de la prensa, ya disponían de esta función desde tiempo atrás y han sido un buen referente.

Esta prestación, muy positiva para el usuario, la han implementado algunas hemerotecas que tienen una página de consulta de nivel aceptable, con opciones de búsqueda simple y avanzada (*Abc*, *Marca*), mientras que en otros casos lo que se ha producido es la sustitución de la búsqueda avanzada por esta función de filtros en la página de resultados (*El país*). En este último caso, las nuevas aportaciones no pueden sustituir plenamente las que se pierden: los filtros son siempre a partir de la consulta por lo que el usuario no

puede saber todas las opciones que hay disponibles antes de realizar la misma.

6.5. Productos documentales en la sección Hemeroteca

Otra tendencia destacable es la transformación de los apartados dedicados en los diarios a sus hemerotecas en secciones amplias que van más allá de las funciones de búsqueda por interrogación o de acceso por navegación cronológica, conteniendo diversos productos y servicios documentales.

Los periódicos que han impulsado con mayor claridad estos cambios son los que poseen hemerotecas centenarias, que han creado secciones muy similares (*Abc*, *La vanguardia*, *Mundo deportivo*), junto a otros como *El mundo* ("*Documenta*", un servicio sólo para suscriptores) o *El país* (que ya tenía "*A fondo*").

Estos contenidos que se ofrecen ahora en la secciones de Hemeroteca pueden ser variados productos documentales (Guallar, 2011), llevados a cabo en buena parte por el servicio de documentación de estos medios, como efemérides, cronologías, perfiles biográficos, o reportajes históricos en los que se utiliza el fondo de hemeroteca; y servicios variados, desde comerciales (como *La tienda de la Hemeroteca de La Vanguardia*) a los de difusión en redes sociales, como

Se tiende a una excesiva simplificación de la consulta, lo que provoca un empobrecimiento de estos sistemas desde el punto de vista documental

La Hemeroteca deja así de ser solamente un servicio en el cual el usuario debe introducir uno o varios términos de consulta en el buscador, marcar unas casillas para acotar la consulta o navegar por ediciones cronológicas. Pasa a ser una nueva sección del diario en la que el fondo retrospectivo llega hasta el lector de diversas maneras, además de permitirle que busque en el mismo. Por ejemplo, mostrándole los hechos más relevantes que sucedieron tal día como hoy y en la forma en que se publicaron hace 25, 50 o incluso 100 años.

7. Notas

1. Los resultados se muestran en detalle en:

Guallar, Javier. *Las hemerotecas de la prensa digital. Análisis de diarios españoles*. Tesis doctoral. Univ. de Barcelona,



Figura 6. Productos documentales en la hemeroteca de La Vanguardia

Dept. de Biblioteconomia i Documentació [en proceso de publicación en TDX, <http://www.tdx.cat>].

Register for free at <https://www.scipedia.com> to download the version without the watermark

de investigación consolidado "Cultura y contenidos digitales" 2009-SGR-177) y forma parte de los proyectos financiados por el Ministerio de Economía y Competitividad, refs.: CSO2012-39518-C04-01 y CSO2012-39518-C04-02.

Los productos documentales en la sección de hemeroteca permiten que el fondo retrospectivo llegue al lector de diversas maneras

8. Bibliografía

Abadal, Ernest. "Elementos para la evaluación de interfaces de consulta de bases de datos". *El profesional de la información*, 2002, v. 11, n. 5, pp. 349-360. <http://www.elprofesionaldelainformacion.com/contenidos/2002/septiembre/3.pdf>

Abadal, Ernest; Guallar, Javier. "Les hemeroteques digitals de la premsa catalana: anàlisi dels diaris catalans de més difusió". En: *11es Jornades catalanes d'informació i documentació*. Barcelona: Cobdc, 2008, pp. 149-152. <http://hdl.handle.net/10760/11696>

Abadal, Ernest; Guallar, Javier. *Prensa digital y bibliotecas*. Gijón: Trea, 2010. 176 pp. ISBN: 978 8494462

Camacho, Idoia; Agirreazaldeg, Teresa; Ronco, Milagros; Peñafiel, Carmen. "¿Evolución o revolución?: Adaptación de los servicios de documentación a la era digital". En: *II Congreso intl. de la Asociación Española de Investigadores de la Comunicación (AEIC)*, Málaga, 2010.

<http://www.aeic2010malaga.org/upload/ok/243.pdf>

Codina, Lluís. "Evaluación de recursos digitales en línea: conceptos, indicadores y métodos". *Revista española de documentación científica*, 2000, v. 23, n. 1 (enero-marzo), pp. 9-44.

<http://redc.revistas.csic.es/index.php/redc/article/view/315/479>

Codina, Lluís. *Metodología de análisis y evaluación de recursos digitales en línea*. Barcelona: UPF, Área de Biblioteconomía y Documentación, Depto. de Periodismo y Comunicación Audiovisual, 2006. 56 pp.

<http://www.lluiscodina.com/metodos/metodos2006.doc>

Guallar, Javier. "Prensa digital en 2006". *Anuario ThinkEpi*, 2007, v. 1, pp.

106-117.

<http://eprints.rclis.org/bitstream/10760/8776/1/AnuarioThinkEPI2007-Guallar-Prensa-digital-en-2006.pdf>

Guallar, Javier. "La documentación en la prensa digital. Nuevas tendencias y perspectivas". En: *III Congreso intl de ciberperiodismo y web 2.0*. Bilbao, Univ. del País Vasco, 9-11 nov. 2011.

<http://eprints.rclis.org/handle/10760/16326>

Guallar, Javier; Abadal, Ernest. "Evaluación de hemerotecas de prensa digital: indicadores y ejemplos de buenas prácticas". *El profesional de la información*, 2009, v. 18, n. 3, pp. 255-269.

<http://hdl.handle.net/10760/13048>

<http://dx.doi.org/10.3145/epi.2009.may.02>

Guallar, Javier; Abadal, Ernest. "The digital press archives of the leading Spanish online newspapers". *Information research*, 2010, v. 15, n. 1.

<http://hdl.handle.net/10760/14439>

Guallar, Javier; Abadal, Ernest; Codina, Lluís. "Sistema de análisis de hemerotecas de prensa digital". *Trípodos*, 2012, n. 31 [en prensa].

Guallar, Javier; Rovira, Cristòfol; Ruiz, Sara. "Multimedialidad en la prensa digital: elementos multimedia y sistemas de recuperación en los principales diarios digitales españoles". *El profesional de la información*, 2010, v. 19, n. 6, pp. 620-629.

<http://eprints.rclis.org/bitstream/10760/15088/1/620-631-Guallar-Rovira-Ruiz.pdf>

<http://dx.doi.org/10.3145/epi.2010.nov.08>

Jiménez-López, Ángels; González-Quesada, Alfons; Fuentes, Maria-Eulàlia. "Las hemerotecas digitales de la prensa en internet". *El profesional de la información*, 2000, v. 9, n. 5, pp. 15-24.

<http://www.elprofesionaldelainformacion.com/contenidos/2000/mayo/2.pdf>

López-Aguirre, José-Luis. "Estudio de servicios y productos documentales en la prensa digital mexicana". *Documentación de las ciencias de la información*, 2009, v. 32, pp. 93-132.

Palacios, Marcos. "La memoria como criterio de valoración de calidad en el ciberperiodismo: algunas consideraciones". *El profesional de la información*, 2009, v. 18, n. 3, pp. 270-277.

<http://dx.doi.org/10.3145/epi.2009.may.03>

Palacios, Marcos; Ribas, Beatriz. "Ferramenta para análise de memória em cibermeios". En: Palacios, Marcos. (coord.). *Ferramentas para análise de qualidade no ciberjornalismo* (vol. 1: Modelos). Portugal: LabCom, 2011, pp. 183-205.

http://www.livroslabcom.ubi.pt/pdfs/20111202-201110_marcos_palacios.pdf

[marcos_palacios.pdf](http://www.livroslabcom.ubi.pt/pdfs/20111202-201110_marcos_palacios.pdf)

Rodríguez-Martínez, Ruth; Codina, Lluís; Pedraza-Jiménez, Rafael. "Cibermedios y web 2.0: modelo de análisis y resultados de aplicación". *El profesional de la información*, 2010, v. 19, n. 1, pp. 35-44.

http://repositori.upf.edu/bitstream/handle/10230/13140/Pedraza_epi_1.pdf?sequence=1

<http://dx.doi.org/10.3145/epi.2010.ene.05>

Rodríguez-Martínez, Ruth; Codina, Lluís; Pedraza-Jiménez, Rafael. "Indicadores para la evaluación de la calidad en cibermedios: análisis de la interacción y de la adopción de la Web 2.0". *Revista española de documentación científica*, 2012, v. 35, n. 1, pp. 61-93.

http://www.digidocweb.net/cibermediosWeb20_2012.pdf

Rubio-Lacoba, María; Blanco-García, Juan-Carlos. "Mejor que un buscador, un encontrador". *Documentación de las ciencias de la información*, n. 33, pp. 273-287.

<http://hdl.handle.net/10760/16306>

Zamith, Fernando. "A methodological proposal to analyze the news websites use of the potentialities of the internet". En: *IX Intl symposium on online journalism*, Texas, 2008.

<http://online.journalism.utexas.edu/2008/papers/Zamith.pdf>

Próximos temas centrales

Enero 2013	Soportes digitales
Marzo 2013	Educación y biblioteca
Mayo 2013	Bibliotecas y documentación de museos
Julio 2013	Economía de la información
Septiembre 2013	Gestión de contenidos
Noviembre 2013	Formación y aprendizaje
Enero 2014	Políticas de información

Los interesados por favor consulten detenidamente la Información para autores:

<http://www.elprofesionaldelainformacion.com/autores.html>

y luego envíen sus artículos a través del gestor de manuscritos OJS de la plataforma del Recyt:

<http://recyt.fecyt.es/index.php/EPI/index>



Creación de Archivos y Bibliotecas Virtuales

Desde la **digitalización** de materiales bibliográficos hasta la asignación de **metadatos** y su **implementación** en la red, conforme a la **normativa internacional**.

Productos para crear Bibliotecas Digitales y Virtuales

DIGIBIB 7.0

Solución avanzada para la creación de Bibliotecas Digitales y la Gestión Bibliotecaria Multilingüe.

DIGIARCH 2.0

Sistema digital de descripción y gestión archivística. Descripción en ISAD(G) y EAD 2.0.

DIGIDIR 2.1

Directorio para Archivos, Bibliotecas y Museos con generación automática de estadísticas y sistemas de información geográfica (GIS).

OASIs-PMH 2.0

Sistema integrado de recolección de diversos esquemas de metadatos:

- DCMI sin cualificar • MARC 21 • EAD
- SWAP • mod_OAI • Linked Open Data

ADAPTACIÓN A EUROPEANA (FASE DANUBIO)

Implementación del esquema **ESE 3.4**

(Europeana Semantic Elements)

y **EDM 5.2.3** (Europeana Data Model)

Adaptado a la Agenda Digital Europea 2020.

DIGITALIZACIÓN AVANZADA

Con asignación dinámica de metadatos.

- Recolección en la Web para Entidades e Instituciones de Memoria en **OAI-PMH** y **Dublin Core** cualificado con **ESE 3.4**
- Consultoría y mappings a **EDM 5.2.3** (Europeana Data Model)
- Bibliotecas digitales que permiten la creación, recuperación y recolección de metadatos (**MARCXML**, **DCMI** y **RDF** y **RDFs**)
- Archivos Web que facilitan la creación, recuperación y recolección de metadatos (**EAD 2.0** y **EAC 2010**)
- Implementación de la **Europeana OpenSearch API**
- Adaptación del repositorio **OAI** para la transmisión de instancias **RDF** según **ORE**
- Repositorios Institucionales **DIGIPRESV** para Preservación Digital a largo plazo mediante **PREMIS 2.2** y **OAIS ISO 14721:2012**
- Intercambio de metadatos en **METS 1.9.1** (diferentes Profiles) integrando todos los esquemas de metadatos
- Creación de **METSRights** para el control de los derechos de autor
- Reconocimiento Óptico de Caracteres **OCR** y generación dinámica de **ALTO 2.0** (Analyzed Layout and Text Object)
- Generación e integración de registros **SKOS** mediante **MARC 21(Up.15)/RDA**
- Enriquecimiento semántico con **LEM**, **VIAF**, **DBpedia** y otros *datasets*
- Adaptación de **DIGIBIB** a **Linked Open Data**
- Creación de eBooks o libros digitales en formatos: **ePub** y **mobi**

ORACLE
PARTNERNETWORK



Validación en el Data Providers de la Open Archives Initiative.
Genera un Sitemap para Google.



ISO 9001
BUREAU VERITAS
Certification
N° ES042816-1

